Proof for Reformulation of Chance-Constrained Optimization under Decision-Dependent Uncertainty

Ning Qi, Pierre Pinson, Mads R. Almassalkhi, Yingrui Zhuang, Yifan Su, Feng Liu

I. Proof of Deterministic Reformulation of Chance-constraint under DDU

Firstly, chance-constraints (1a), (1b) in the GES model admit a compact form (1c). Hereby, $a_i(\boldsymbol{x})$,$b_i(\mathbf{x})$ are affine functions of decisions $\boldsymbol{x}$. $\boldsymbol{\xi}(\boldsymbol{x})$ defines decision-dependent uncertainty (DDU) with decision-dependent mean $\boldsymbol{\mu}(\boldsymbol{x})$ and decision-independent covariance $\boldsymbol{\Sigma}$. Machine learning methods can be used to estimate the structure and distribution of DDU [1], [2] from the observed demand response performance in CAISO [3]. Deterministic reformulation of (1a),(1b) is shown in (1d). While, DDU causes the inverse cumulative distribution function F^{-1} to be coupled with decisions $\boldsymbol{x}$, making it impossible to know its numerical value before optimization. Hence, we provide three effective approaches to address this issue based on the acquired information of DDU.

$$\mathbb{P}(\underline{SoC}_{i,t}^{\mathrm{DDU}} \le SoC_{i,t}^{\mathrm{RT}}) \ge 1-\epsilon \tag{1a}$$

$$\mathbb{P}(P_{\mathrm{d},i,t}^{\mathrm{RT}} \le \overline{P}_{\mathrm{d},i,t}) \ge 1-\epsilon \tag{1b}$$

$$\mathbb{P}\Big(a_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\xi}(\boldsymbol{x}) \le b_i(\boldsymbol{x})\Big) \ge 1-\epsilon \tag{1c}$$

$$a_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}(\boldsymbol{x}) + b_i(\boldsymbol{x}) + F_{\boldsymbol{x}}^{-1}(1-\epsilon)\sqrt{a_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}a_i(\boldsymbol{x})} \le 0 \tag{1d}$$

(i) Robust Approximation: for DDU with general but ambiguous distribution, a robust approximation is introduced. Hereby, generalizations of the Cantelli's inequality can be used to estimate the best probability bound (i.e., the maximum value of $F_{\boldsymbol{x}}^{-1}$). The maximum values of $F_{\boldsymbol{x}}^{-1}$ for six widely used distributions are derived and listed in Table I and the visualization is shown in Fig 1. These can be readily employed in any CCO-DDUs problems without complete knowledge of DDUs distribution. The supporting proofs are provided in Appendix B. It is observed that the value decreases with increasing security levels. Besides, the first 4 approximation types listed in Table II relies on less information about the type of distribution at hand. Consequently, they lead to more conservative approximations (i.e., as a higher value for $F_{\boldsymbol{x}}^{-1}$), which will further lead to higher security levels and tighter bounds. Since we do not know the exact distribution of DDUs in advance, but at least we can obtain the approximate shape of the distribution (e.g., unimodal or symmetric, etc.) through some live measurements or prior knowledge. For instance, if the unknown distribution is a Beta-like distribution, the approximation type for a unimodal distrtribution (3rd entry in Table II) can be used to replace $F_{\boldsymbol{x}}^{-1}$ in the reformulation, and chance-constrained problem under DDU are then reduced to chance-constrained problem under DIU. This eventually yields a robust reformulation that is less conservative than using that without any distributional assumption (first entry in Table II). Noted that robust approximations can generate over-conservative solutions to chance-constrained problems under DDU, but at least they guarantee that the practical performance of the response lies within the required security level. And, it is especially applicable for the black-start of system without sufficient historical data of GES.

TABLE I
Approximation of Widely Used Normalized Inverse Cumulative Distribution

Type & Shape	$\overline{F}_{\boldsymbol{x}}^{-1}(1-\epsilon)$	ϵ
1) No distribution assumption (NA)	$\sqrt{(1-\epsilon)/\epsilon}$	$0<\epsilon\le 1$
2) Symmetric distribution (S)	$\sqrt{1/2\epsilon}$	$0<\epsilon\le 1/2$
	0	$1/2<\epsilon\le 1$
3) Unimodal distribution (U)	$\sqrt{(4-9\epsilon)/9\epsilon}$	$0<\epsilon\le 1/6$
	$\sqrt{(3-3\epsilon)/(1+3\epsilon)}$	$1/6<\epsilon\le 1$
4) Symmetric & unimodal distribution (SU)	$\sqrt{2/9\epsilon}$	$0<\epsilon\le 1/6$
	$\sqrt{3}(1-2\epsilon)$	$1/6<\epsilon\le 1/2$
	0	$1/2<\epsilon\le 1$
5) Student's t distribution (ST)	$t_{\nu,\sigma}^{-1}(1-\epsilon)$	$0<\epsilon\le 1$
6) Normal distribution (N)	$\Phi^{-1}(1-\epsilon)$	$0<\epsilon\le 1$

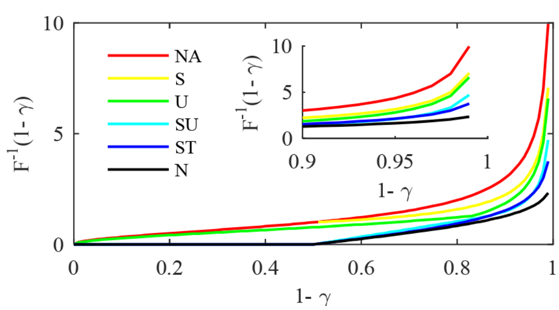

Fig. 1. Visualization of inverse CDF with six types of distribution.

The proof of Table 1 is further provided. Hereby, we write F the CDF function, $\mathbb{P}$ the PDF function, $k\ge 0$ a constant, and ξ as the probabilistic parameter with zero mean and unit variance under the chosen distribution. Different versions of Cantelli's inequality [4] are used to obtain the following results.

1) Classical Cantelli inequality can be used without distribution assumption of DDUs and infers the following conclusion.

$$F(k) = 1 - \underset{P\in NA}{sup}\ \mathbb{P}[\xi \ge k] = k^2/1+k^2 \tag{2a}$$

$$F^{-1}(1-\epsilon) = \sqrt{(1-\epsilon)/\epsilon} \tag{2b}$$

2) Chebyshev's inequality can be used with symmetric

distribution of DDUs and infers the following conclusion.

$$F(k) = 1 - \sup_{P \in S} \mathbb{P}[\xi \geq k] = 1 - \frac{1}{2} \sup_{P \in S} \mathbb{P}[|\xi| \geq k] = 1 - \frac{1}{2k^2} \quad \text{(3a)}$$

$$F^{-1}(1-\epsilon) = \sqrt{1/2\epsilon} \quad \text{(3b)}$$

3) VySoChanskij–Petunin inequality can be used with unimodal distribution of DDUs and infers the following conclusion.

$$\begin{aligned} F(k) &= 1 - \sup_{P \in U} \mathbb{P}[\xi \geq k] \\ &= \begin{cases} 1 - 4/(9k^2+9) & k \geq \sqrt{5/3} \\ 1-(3-k^2)/(3+3k^2) & 0 \leq k \leq \sqrt{5/3} \end{cases} \end{aligned} \quad \text{(4a)}$$

$$F^{-1}(1-\epsilon) = \begin{cases} \sqrt{2/9\epsilon} & 0 < \epsilon \leq 1/6 \\ \sqrt{3}(1-2\epsilon) & 1/6 < \epsilon \leq 1/2 \end{cases} \quad \text{(4b)}$$

4) Gauss's inequality can be used for symmetric & unimodal distribution of DDUs and infers the following conclusion.

$$\begin{aligned} F(k) &= 1 - \sup_{P \in SU} \mathbb{P}[\xi \geq k] = 1 - \frac{1}{2} \sup_{P \in U} \mathbb{P}[|\xi| \geq k] \\ &= \begin{cases} 1 - 2/9k^2 & k \geq 2/\sqrt{3} \\ 1/2 + k/2\sqrt{3} & 0 \leq k \leq 2/\sqrt{3} \end{cases} \end{aligned} \quad \text{(5a)}$$

$$F^{-1}(1-\epsilon) = \begin{cases} \sqrt{2/9\epsilon} & 0 < \epsilon \leq 1/6 \\ \sqrt{3}(1-2\epsilon) & 1/6 < \epsilon \leq 1/2 \end{cases} \quad \text{(5b)}$$

5-6) For student's t and normal distribution of DDUs, the normalized CDFs $t^{-1}_{\nu,\sigma}(1-\epsilon)$ and $\Phi^{-1}(1-\epsilon)$ can be used without introducing approximation errors.

(ii) Iterative Optimization: We propose an iterative algorithm in Algorithm 1 for more precise structure (known function and distribution of g and h), if sufficient live measurement/data about GES are provided. First, the robust reformulation method is used as the starting point of the inversed CDF, which generates the most conservative result $\boldsymbol{x}^{(0)}$. Afterward, the iteration begins with the updated value of DDU $\boldsymbol{\xi}(\boldsymbol{x}^{(k)})$ and $F^{-1}_{\boldsymbol{x}^{(k)}}(1-\epsilon)$ to obtain the updated decisions $\boldsymbol{x}^{(k)}$. The update of $F^{-1}_{\boldsymbol{x}^{(k)}}(1-\epsilon)$ is computed via Monte Carlo Sampling of the updated distribution. he iterations stop when the the convergence criterion is met. The convergence of the iterative algorithm is guaranteed by the convexity of chance constrained optimization under DDU, exactly the convexity of the mean function of DDU since the variance is decision-independent here. The iterative algorithm is illustrated in Eq. (6).

$$\overline{F}^{-1}_{\boldsymbol{x}}(1-\epsilon) \rightarrow \boldsymbol{x}^{(0)} \rightarrow \boldsymbol{\xi}(\boldsymbol{x}^{(0)}) \rightarrow F^{-1}_{\boldsymbol{x}^{(0)}}(1-\epsilon) \rightarrow \boldsymbol{x}^{(1)} \rightarrow \cdots \boldsymbol{x}^{(k)} \quad \text{(6)}$$

Next we provide the convexity and convergence proof, we adopt a convex structure of DDU in this paper, as shown in (7) below,

$$g = \begin{cases} (\overline{SoC}^{\mathrm{PY}}_{i,t} - \overline{SoC}^{\mathrm{DIU}}_{i,t})\mathcal{N}(\mu_{g^{\mathrm{U}}}, \sigma_g) + \overline{SoC}^{\mathrm{DIU}}_{i,t} \\ (\underline{SoC}^{\mathrm{PY}}_{i,t} - \underline{SoC}^{\mathrm{DIU}}_{i,t})\mathcal{N}(\mu_{g^{\mathrm{L}}}, \sigma_g) + \underline{SoC}^{\mathrm{DIU}}_{i,t} \end{cases} \quad \text{(7a)}$$

$$h = \begin{cases} (\overline{SoC}^{\mathrm{C}}_{i,t} - Q_{g^{\mathrm{U}}})\mathcal{LN}(\mu_{h^{\mathrm{U}}}, \sigma_h) + Q_{g^{\mathrm{U}}} \\ (\underline{SoC}^{\mathrm{C}}_{i,t} - Q_{g^{\mathrm{L}}})\mathcal{LN}(\mu_{h^{\mathrm{L}}}, \sigma_h) + Q_{g^{\mathrm{L}}} \end{cases} \quad \text{(7b)}$$

$$\mu_{g^{\mathrm{U/L}}} = c^{\mathrm{S}}_{\mathrm{c/d},i,t}/\overline{c}^{\mathrm{S}}, \; \mu_{h^{\mathrm{U/L}}} = \beta^{\mathrm{U/L}}_i RD_{i,t}, \quad \text{(7c)}$$

According to the convexity condition of CCO and reformulations above, chance-constrained optimization under DDU is only guaranteed to be convex under the condition (i)-(ii). The convergence of the iterative algorithm is guaranteed when the convexity condition is satisfied [5].

(i) $\mu_{\underline{SoC}_{i,t}}$ and $F^{-1}_{\underline{SoC}_{i,t}}(1-\epsilon)\sigma_{\underline{SoC}_{i,t}}$ are convex function of decision variables $\boldsymbol{x}$.

(ii) $-\mu_{\overline{SoC}_{i,t}}$ and $F^{-1}_{\overline{SoC}_{i,t}}(1-\epsilon)\sigma_{\overline{SoC}_{i,t}}$ are convex function of decision variables $\boldsymbol{x}$.

For DDU designed in (7), the inside functions are given:

$$\mu_{\overline{SoC}_{i,t}} = (\overline{SoC}^{\mathrm{B}}_{i,t} - Q_{g^{\mathrm{U}}})\beta^{\mathrm{U}}_i RD_{i,t} + Q_{g^{\mathrm{U}}} \quad \text{(8a)}$$

$$\mu_{\underline{SoC}_{i,t}} = (\underline{SoC}^{\mathrm{B}}_{i,t} - Q_{g^{\mathrm{D}}})\beta^{\mathrm{D}}_i RD_{i,t} + Q_{g^{\mathrm{D}}} \quad \text{(8b)}$$

$$F^{-1}_{\overline{SoC}_{i,t}}(1-\epsilon)\sigma_{\overline{SoC}_{i,t}} = (Q_{g^{\mathrm{U}}} - \overline{SoC}^{\mathrm{B}}_{i,t})F^{-1}_{h^{\mathrm{U}}}(1-\epsilon, \boldsymbol{y})\sigma_{h^{\mathrm{U}}} \quad \text{(8c)}$$

$$F^{-1}_{\underline{SoC}_{i,t}}(1-\epsilon)\sigma_{\underline{SoC}_{i,t}} = (Q_{g^{\mathrm{D}}} - \underline{SoC}^{\mathrm{B}}_{i,t})F^{-1}_{h^{\mathrm{D}}}(1-\epsilon, \boldsymbol{y})\sigma_{h^{\mathrm{D}}} \quad \text{(8d)}$$

Thus, the convexity conditions are further simplified as:
(a) $RD_{i,t}$ is a convex function of $\boldsymbol{x}$.
(b) $F^{-1}_{\boldsymbol{x}}(1-\epsilon)$ is a convex function of $\boldsymbol{x}$.

The convex function of response discomfort guarantees the convexity condition (a). And since the variance is decision-independent, the convexity of $F^{-1}_{\boldsymbol{x}}(1-\epsilon)$ is equivalent to the convexity of $F^{-1}_{\boldsymbol{\mu}}(1-\epsilon)$. For lognormal distributions, $F^{-1}_{\boldsymbol{\mu}}(1-\epsilon) = \exp(\mu + \sqrt{2\sigma^2}\,\mathrm{erf}^{-1}(1-2\epsilon))$, which guarantees the convexity condition (b). While, for other complex distributions (e.g., Beta), there is no explicit expression for the inverse CDF, and numerical simulations in Fig. 3 shows that it can not guarantee convexity overall. There exists, however, a convex region which contains the iterations using Beta distribution. Thus, global optimality can be verified for this convex region and corresponding constraints can be added to limit response discomfort (μ) of GES units within that region.

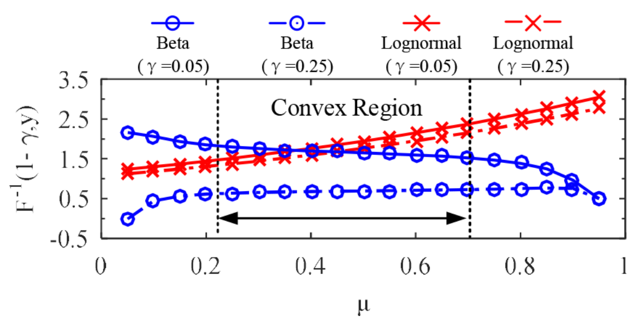

Fig. 2. Numerical test of convexity

(iii) Data-Driven Approach: the robust approximation method is conservative, but if DDU can be observed in real-time operation, the chance-constrained optimization under DDU will be simplified with distributionally robust chance-constrained optimization under DIU. Hereby, a data-driven reformulation is provided in (9), $\boldsymbol{r_x}$ is the radius of the observed DDU, $\boldsymbol{y}$ is the auxiliary decision matrix, constant p and observed samples K should guarantee that: $p \geq 2$, $K > (2+\sqrt{2\ln(4/\epsilon)})^p$.

$$\begin{cases} a_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}(\boldsymbol{x}) + b_i(\boldsymbol{x}) + \psi_K \|\boldsymbol{r}(\boldsymbol{x})\|_1 + \pi_K\sqrt{1/\epsilon - 1}\,\|\boldsymbol{y}\|_2 \leq 0 \\ \sqrt{a_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma} a_i(\boldsymbol{x})} \leq y_1, \quad \sqrt{2\psi_K}\,\|\boldsymbol{r}(\boldsymbol{x})\|_1 \leq y_2 \\ \psi_K = K^{(1/p-1/2)}, \quad \pi_K = \left(1 - \frac{4}{\epsilon}\exp(-(K^{1/p}-2)^2/2)\right)^{-1/2} \end{cases} \quad \text{(9)}$$

REFERENCES

[1] J. Ruan, G. Liang, J. Zhao, S. Lei, B. He, J. Qiu, and Z. Y. Dong, "Graph deep-learning-based retail dynamic pricing for demand response," *IEEE Trans. on Smart Grid*, vol. 14, no. 6, pp. 4385–4397, 2023.

[2] L. Marrero, D. Sbárbaro, and L. García-Santander, "Online demand response characterization based on variability in customer behavior," *Journal of Modern Power Systems and Clean Energy*, 2024.

[3] CAISO, "Demand response issues and performance." [Online]. Available: http://www.caiso.com/participate/Pages/Load/Default.aspx.

[4] E. Roos, R. Brekelmans, and et al., "Tight tail probability bounds for distribution-free decision making," *European Journal of Operational Research*, vol. 299, no. 3, pp. 931–944, 2022.

[5] C. T. Kelley, *Iterative methods for optimization*. SIAM, 1999.